PCT/GB 98 / 02630

09/486882

The Patent Office

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP9 1RH

PRIORITY DOCUMENT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

I also certify that the attached copy of the request for grant of a Patent (Form 1/77) bears an amendment, effected by this office, following a request by the applicant and agreed to by the Comptroller-General.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed Andrew Gersey

Dated 5th October 1998

Patents Form 1/77

Patents Act 1977
(Rule 16)

The

9718455.0

1/77

THE PATENT OFFICE
-2 SEP 1997
RECEIVED BY POST

2 SEP 1997

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference DMPAT004

2. Patent application number *(The Patent Office will fill in this part)*
02SEP97 E299641-1 C25514
P01/7700 25.00 - 9718455.0

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*
DUNCAN PATRICK McGREGOR
6 BALCAIRN COTTAGES
OLDMELDRUM
ABERDEENSHIRE AB51 0EU

Patents ADP number *(if you know it)*

If the applicant is a corporate body, give the country/state of its incorporation GB 7711900002.

4. Title of the invention ~~CHIMERIC BINDING PEPTIDE LIBRARY SCREENING METHOD~~ D.L.
CHIMERIC BINDING PEPTIDE LIBRARY SCREENING METHOD

5. Name of your agent *(if you have one)*

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*
~~6 BALCAIRN COTTAGES~~
~~OLDMELDRUM~~
~~ABERDEENSHIRE~~
~~AB51 0EU~~
Murgitroyd & Company
373 Scotland Street
Glasgow
G5 8QA

Patents ADP number *(if you know it)*
P5177 2/9/98 4P.

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*
a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	0
Description	13
Claim(s)	1
Abstract	1
Drawing(s)	3+M

10. If you are also filing any of the following, state how many against each item.

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	

11. I/We request the grant of a patent on the basis of this application.

Signature Duncan McGregor Date 31/8/97

12. Name and daytime telephone number of person to contact in the United Kingdom DUNCAN McGREGOR 01224 712751 x3120

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.

b) Write your answers in capital letters using black ink or you may type them.

c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.

d) If you have answered 'Yes' Patents Form 7/77 will need to be filed.

e) Once you have filled in the form you must remember to sign and date it.

f) For details of the fee and ways to pay please contact the Patent Office.

Chimeric binding peptide library screening method

The present invention relates generally to recombinant DNA technology and, more particularly, to methods for screening DNA libraries for DNA sequences that encode biological binding molecules.

Background

Isolating an unknown gene which encodes a desired peptide from a recombinant DNA library can be a difficult task. The use of hybridisation probes may facilitate the process, but their use is generally dependent on knowing at least a portion of the sequence of the gene which encodes the protein. When the sequence is not known, DNA libraries can be expressed in an expression vector, and antibodies have been used to screen plaques or colonies for the desired protein antigen. This procedure has been useful in screening small libraries, but rarely occurring sequences which are represented in less than about 1 in 10^5 clones, as is the case with rarely occurring cDNA molecules or synthetic peptides, can be easily missed, making screening libraries larger than 10^6 clones at best laborious and difficult. Screening larger libraries has required the development of methods designed to address the isolation of rarely occurring sequences:

(a.) Phage display methods. DNA libraries fused to the N-terminal end of filamentous bacteriophage pIII and pVIII coat proteins have been expressed from an expression vector resulting in the display of foreign peptides on the surface of the phage particle with the DNA encoding the fusion protein packaged in the phage particle (Smith G. P., 1985, Science 228: 1315-1317). The expression vector can be the bacteriophage genome itself, or a phagemid vector, into which a bacteriophage coat protein has been cloned. In the latter case, the host bacterium, containing the phagemid vector, must be co-infected with autonomously replicating bacteriophage, termed helper phage, to provide the full complement of proteins necessary to produce mature phage particles. The helper phage normally has a genetic defect in the origin of replication which results in the preferential packaging of the phagemid genome. Expression of the fusion protein following helper phage infection, allows incorporation of both fusion protein and wild type coat protein into the phage particle during assembly. Libraries of fusion proteins incorporated into phage, can then be selected for binding members against targets of interest (ligands). Bound phage can then be allowed to reinfect *Escherichia coli* (*E. coli*) bacteria and then amplified and the selection repeated, resulting in the enrichment of binding members (Parmley, S. F., & Smith, G. P. 1988., Gene 73: 305-318; Barrett R. W. *et al.*, 1992, Analytical Biochemistry 204: 357-364 Williamson *et al.*, Proc. Natl. Acad. Sci. USA, 90: 4141-4145; Marks *et al.*, 1991, J. Mol. Biol. 222: 581-597).

Several disclosures have been made for this method. For example, in the teaching of US5,403,484, at least a functional portion of a viral coat protein is required, to cause the display of the chimeric binding protein or a processed form thereof on the outer surface of the virus. In addition, US5,571,698 claims a method for obtaining a nucleic acid encoding a binding protein, a key component of which comprises preparing a population of amplifiable genetic packages which have a genetically determined outer surface protein. This is further defined as an outer surface transport signal for obtaining the display of the potential binding domain on the outer surface of the genetic package. The genetic packages being selected from the group consisting of cells, spores and viruses. The method teaches that when the genetic package is a bacterial cell, the outer

surface transport signal is derived from a bacterial outer surface protein, and when the genetic package is filamentous bacteriophage, the outer surface transport signal is provided by the gene pIII (minor coat protein) or pVIII (major coat protein) of the filamentous phage.

In other disclosures (WO92/01047 and WO92/20791), methods are described for producing multimeric specific binding pairs, by expressing a first polypeptide chain fused to a component of a secreted replicable genetic display package (RGDP) which displays a polypeptide at the surface of the package, and expressing a second polypeptide chain of the multimer, and allowing the two chains to come together as part of the RGDP. It is also disclosed that the component of the RGDP should be a viral coat protein. The example of viral coat proteins disclosed is the gene pIII protein.

(b.) LacI fusion plasmid display. This method is based on the DNA binding ability of the lac repressor. Libraries of random peptides are fused to the lacI repressor protein, normally to the C-terminal end, through expression from a plasmid vector carrying the fusion gene. Linkage of the LacI-peptide fusion to its encoding DNA occurs via the lacO sequences on the plasmid, forming a stable peptide-LacI-peptide complex. These complexes are released from their host bacteria by cell lysis, and peptides of interest isolated by affinity purification on an immobilised target. The plasmids thus isolated can then be reintroduced into *E. coli* by electroporation to amplify the selected population for additional rounds of screening (Cull, M. G. et al. 1992. Proc. Natl. Acad. Sci. U.S.A. 89:1865-1869).

US5498530 discloses a method for constructing a library of random peptides fused to a DNA binding protein in appropriate host cells and culturing the host cells under conditions suitable for expression of the fusion proteins intra-cellularly, in the cytoplasm of the host cells. This method also teaches that the random peptide is located at the carboxy terminus of the fusion protein and that the fusion protein-DNA complex be released from the host cell by cell lysis. No method is provided for the protection of the DNA from degradation once released from the lysed cell. Several DNA binding proteins are claimed but no examples are shown except for lacI.

There remains a need for methods of constructing peptide libraries in addition to the methods described above. For instance, the above methods do not provide secreted peptides with a free carboxy terminus. The invention described herein, describes an alternative method for isolating peptides of interest from libraries with significant advantages over the prior art methods.

Summary of the invention

Methods are provided for screening a DNA library for a nucleotide sequence which encodes a peptide of interest. The methods generally involve physically linking the peptide, in a biologically active form (usually having a binding activity) to the specific library member DNA sequence encoding that peptide, through fusion to a DNA binding domain (DBD). This results in the production of a bifunctional chimeric protein with a DNA binding domain and a DNA library member domain with a function of interest, hereafter referred to as a chimeric protein. The chimeric protein binds to a specific nucleotide target sequence present in the DNA containing the nucleotide sequence encoding the chimeric protein. This DNA-chimeric protein complex is then incorporated within a peptide display carrier package (PDCP), protecting that DNA from subsequent

degradation, while displaying the DNA library member peptide of interest on the outer surface of the peptide display carrier package (PDCP). Said PDCP is composed two distinct elements:

a. A DNA-chimeric protein complex. This links the displayed peptide to the library member DNA sequence encoding that peptide using a specific DNA binding domain. The nucleotide sequence encoding the chimeric protein, and the DBD specific nucleotide target sequence, must be present on a segment of DNA which can be incorporated into the PDCP.

b. A protective coat supplied by a replicable carrier package capable of independent existence. The replicable carrier package supplies a protective coat for the DNA-chimeric protein complex which maybe composed of a biological material such as protein or lipid, but is not required for linking the displayed peptide to the DNA encoding that peptide. Said carrier package must allow the display of the DNA library member peptide component of the chimeric protein on the outer surface of the protective coat. The carrier package generally also provides the mechanism for releasing the intact PDCP from host cells when so required. By way of example, when the replicable carrier package is a bacteriophage, a protein coat surrounds the DNA-chimeric protein complex to form the PDCP, which is then extruded from the host bacterial cell.

Surprisingly, the invention described herein, demonstrates that peptides fused to a DNA binding domain can be displayed externally, through a bacteriophage carrier package protein coat, while still bound to the DNA encoding the displayed peptide. Also disclosed are methods for constructing and screening libraries of PDCP particles, displaying many different peptides, allowing the isolation and identification of particular peptides by means of affinity techniques relying on the binding activity of the peptide of interest. The resulting DNA sequences can therefore be more readily identified, re-cloned and expressed by using this invention. The current invention differs from the prior art teaching of the previous disclosures US5,403,484 or US5,571,698, as the invention does not require outer surface transport signals, or functional portions of viral coat proteins, to enable the display of chimeric binding proteins on the outer surface of the viral particle or genetic package. The current invention differs from the teaching of WO92/01047 and WO92/20791, as no component of a secreted replicable genetic display package, or viral coat protein is required, to enable display on the outer surface of the viral particle. The current invention differs from the teaching of US5498530, as it enables the display of chimeric proteins, linked to the DNA encoding the chimeric protein, extra-cellularly, not in the cytoplasm of a host cell. In the current invention the chimeric proteins are presented on the outer surface of a peptide display carrier package (PDCP) which protects the DNA encoding the chimeric protein, and does not require cell lysis to obtain access to the chimeric protein-DNA complex. Finally, the current invention does not use the lacI DNA binding protein to form the chimeric protein-DNA complex.

Description of the invention

In a preferred embodiment of the invention, DNA binding proteins used to form the DNA-chimeric protein complex of the PDCP, include members of the nuclear steroid receptor family of proteins, such as the oestrogen receptor and the progesterone receptor, amongst others. In a more preferred embodiment the DNA binding proteins are limited to the DNA binding domains of these proteins. These domains can recognise specific DNA sequences, termed hormone response elements (HRE), which can be bound as both double and single-stranded DNA. The DNA binding domain of such proteins is preferred, using the oestrogen receptor by way of example not exclusion, for a number of reasons:

(a) The oestrogen receptor is a large multifunctional polypeptide of 595 amino acids which functions in the cytoplasm and nucleus of eukaryotic cells (Green et al., 1986, Science 231: 1150-1154). A minimal high affinity DNA binding domain (DBD) has been defined between amino acids 176 and 282 (Mader et al., 1993, Nucleic Acids Res. 21: 1125-1132). The functioning of this domain i.e. DNA binding is not inhibited by the presence of non-DNA binding domains at both the N and C terminal ends of this domain, in the full length protein.

(b) The oestrogen receptor DNA binding domain fragment (amino acids 176-282) has been expressed in *E. coli* and shown to bind to the specific double stranded DNA oestrogen receptor target HRE nucleotide sequence, as a dimer with a similar affinity (0.5nM) to the parent molecule (Murdoch et al. 1990, Biochemistry 29: 8377-8385; Mader et al., 1993, Nucleic Acids Research 21: 1125-1132). DBD dimerization on the surface of the PDCP should result in two peptides displayed per particle. This bivalent display can aid in the isolation of low affinity peptides and peptides that are required to form a bivalent conformation in order to bind to a particular target, or activate a target receptor.

(c) The oestrogen receptor is capable of binding to its 38 base pair target HRE sequence, consensus sequence:

5′-TCAGGTCAGAGTGACCTGAGCTAAAATAACACATTCAG-3′ ("minus strand")

3′-AGTCCAGTCTCACTGGACTCGATTTTATTGTGTAAGTC-5′ ("plus strand")

with high affinity and specificity, under the salt and pH conditions normally required for selection of binding peptides. Moreover, binding affinity is increased 60-fold for the single-stranded coding, or "plus," strand of the HRE nucleotide sequence over the double stranded form of the specific target nucleotide sequence (Peale et al. 1988, Proc. Natl. Acad. Sci. USA 85: 1038-1042; Lannigan & Notides, 1989, Proc. Natl. Acad. Sci. USA 86: 863-867).

In an embodiment of the invention where the DNA binding component of the peptide display carrier package is the oestrogen receptor, the DNA binding domain contains a minimum sequence of amino acids 176-282 of the oestrogen receptor protein. In addition, the consensus oestrogen receptor target HRE sequence is cloned in such a way that if single stranded DNA can be produced then the coding, or "plus," strand of the oestrogen receptor HRE nucleotide sequence is incorporated into single-stranded DNA. An example of a vector suitable for this purpose is pUC119.

In a preferred embodiment of the invention a peptide display carrier package (PDCP) can be assembled when a bacterial host cell is transformed with a bacteriophage vector, which vector comprises a DNA library member joined to a nucleotide sequence encoding a DNA binding protein, or a functional DBD thereof, which can bind

to a specific target nucleotide sequence. The expression vector also contains the specific target nucleotide sequence that can be bound by the DBD. Expression of this DNA sequence results in the production of a chimeric protein which comprises the DNA library member peptide and the DBD. The host cell is grown under conditions suitable for chimeric protein expression and assembly of the bacteriophage particles, and the association of the chimeric protein with the specific DNA nucleotide sequence in the expression vector. Incorporation of the vector single-stranded DNA-chimeric protein complex into bacteriophage particles results in the assembly of the peptide display carrier package (PDCP), and display of the DNA library member peptide on the outer surface of the PDCP. In this embodiment both of the required elements for producing peptide display carrier packages are contained on the same vector. Incorporation of the DNA-chimeric protein complex into a peptide display carrier package (PDCP) is preferred as DNA degradation is prevented, large numbers of PDCPs' are produced per host cell, and the PDCPs are easily separated from the host cell without recourse to cell lysis.

In a more preferred embodiment, the vector of the previous embodiment is a phagemid vector, an example of which is pUC119, where expression of the chimeric protein is controlled by an inducible promoter. In this embodiment the PDCP can only be assembled following infection of the host cell, with helper phage and then cultivated under conditions suitable for chimeric protein expression and assembly of the bacteriophage particles. In this embodiment the elements of the PDCP are provided by two separate vectors. The phagemid derived PDCP is superior to phagemid derived display packages disclosed in WO92/01047, as in the prior art, a proportion of packages displaying bacteriophage coat protein fusion proteins will contain the helper phage DNA, not the fusion protein DNA sequence. In the current invention, a PDCP can display the chimeric fusion protein, only when the package contains the DBD specific nucleotide target sequence. In most embodiments this sequence will be present on the same DNA segment that encodes the fusion protein. In addition, the prior art acknowledges that when mutant and wild type proteins are co-expressed in the same bacterial cell, the wild type protein is used preferentially. Thus, when the wild type helper phage, phage display system of WO92/01047 is used, both wild type gene pIII and DNA library member peptide-gene pIII chimeric proteins are produced in the same cell. The result of this is that the wild type gene pIII protein is preferentially packaged into bacteriophage particles, over the chimeric protein. In the current invention, there is no competition with wild type bacteriophage coat proteins for packaging.

In a preferred embodiment of the invention the DNA library member peptide is displayed in a location exposed to the external environment of the PDCP, after the PDCP particle has been released from the host cell without recourse to cell lysis, and is accessible for binding to its ligand. In this embodiment the peptide is located at or near the N-terminus or the C-terminus of a DNA binding domain (DBD), of, for example, the oestrogen receptor, and concerns methods for screening a DNA library expressing one or more polypeptide chains that are processed, folded and assembled in the periplasmic space to achieve biological activity. The PDCP is assembled by the following steps:

(a) Construction of N- and C-terminal DBD chimeric protein fusions in a phagemid vector.

(i) When the DNA library member peptide is located at the N-terminus of the DBD, a library of DNA member sequences is cloned into an appropriate location of the expression vector, behind an appropriate promoter and translation sequences and a sequence encoding a signal peptide leader directing transport of the downstream fusion protein to the periplasmic space. In a preferred embodiment the DNA sequence(s) of interest is joined, by a region of DNA encoding a flexible amino acid linker, to the 5'-end of an oestrogen receptor DBD.

(ii) When the DNA library member peptide located is at the C-terminus of the DBD, a library of DNA member sequences is cloned into the expression vector so that the nucleotide sequence coding for the DBD of the oestrogen receptor is upstream of the cloned DNA library member sequences, said DBD being positioned behind an appropriate promoter and translation sequences and a sequence encoding a signal peptide leader directing transport of the downstream fusion protein to the periplasmic space. The DNA sequence(s) of interest is joined, by a region of DNA encoding a flexible amino acid linker, to the 3'-end of an oestrogen receptor DBD DNA sequence.

Located on the expression vector is the specific HRE nucleotide sequence recognised, and bound, by the oestrogen receptor DBD. In order to vary the number of chimeric proteins displayed on each PDCP particle, this sequence can be present as one or more copies in the vector.

(b) Incorporation into the PDCP. Non-lytic helper bacteriophage are allowed to infect host cells containing the expression vector. In a preferred embodiment the bacteriophage include the filamentous phage fd, f1 and M13. in a more preferred embodiment the bacteriophage is M13K07. The protein(s) of interest are expressed and transported to the periplasmic space, and the properly assembled proteins are incorporated into the PDCP particle by virtue of the high affinity interaction of the DBD with the specific target nucleotide sequence present on the phagemid vector DNA which is naturally packaged into phage particles in a single-stranded form. The high affinity interaction between the DBD protein and its specific target nucleotide sequence prevents displacement by bacteriophage coat proteins resulting in the incorporation of the protein(s) of interest onto the surface of the PDCP as it is extruded from the cell.

(c) Selection of the peptide of interest. Particles which display the peptide of interest are then selected from the culture by affinity enrichment techniques. This is accomplished by means of a ligand specific for the protein of interest, such as an antigen if the protein of interest is an antibody. The ligand may be presented on a solid surface such as the surface of an ELISA plate, or in solution. Repeating the affinity selection procedure provides an enrichment of clones encoding the desired sequences, which may then be isolated for sequencing, further cloning and/or expression.

Numerous types of libraries of peptides fused to the DBD can be screened under this embodiment including:

(i) Random peptide sequences encoded by synthetic DNA of variable length.

(ii) Single-chain Fv antibody fragments. These consist of the antibody heavy and light chain variable region domains joined by a flexible linker peptide to create a single-chain antigen binding molecule.

(iii) Random fragments of naturally occurring proteins isolated from a cell population containing an activity of interest.

In another embodiment the invention concerns methods for screening a DNA library whose members require more than chain for activity, as required by, for example, antibody Fab fragments for ligand binding. In this embodiment heavy or light chain antibody DNA is joined to a nucleotide sequence encoding a DNA binding domain of, for example, the oestrogen receptor in a phagemid vector. Typically the antibody DNA library sequences for either the heavy (VH and CH1) or light chain (VL and CL) genes are inserted in the 5' region of the oestrogen receptor DBD DNA, behind an appropriate promoter and translation sequences and a sequence encoding a signal peptide leader directing transport of the downstream fusion protein to the periplasmic space. Thus, a DBD fused to a DNA library member-encoded protein is produced and assembled in to the viral particle after infection with bacteriophage. The second and any subsequent chain(s) are expressed separately either:

(a) from the same phagemid vector containing the DBD and the first polypeptide fusion protein,

or

(b) from a separate region of DNA which may be present in the host cell nucleus, or on a plasmid, phagemid or bacteriophage expression vector that can co-exist, in the same host cell, with the first expression vector, so as to be transported to the periplasm where they assemble with the first chain that is fused to the DBD protein as it exits the cell. Peptide display carrier packages (PDCP) which encode the protein of interest can then be selected by means of a ligand specific for the protein.

In yet another embodiment, the invention concerns screening libraries of bi-functional peptide display carrier packages where two or more activities of interest are displayed on each PDCP. In this embodiment, a first DNA library sequence(s) is inserted next to a first DNA binding domain (DBD) DNA sequence, for example the oestrogen receptor DBD, in an appropriate vector, behind an appropriate promoter and translation sequences and a sequence encoding a signal peptide leader directing transport of this first chimeric protein to the periplasmic space. A second chimeric protein is also produced from the same, or separate, vector by inserting a second DNA library sequence(s) next to a second DBD DNA sequence which is different from the first DBD DNA sequence, for example the progesterone receptor DBD, behind an appropriate promoter and translation sequences and a sequence encoding a signal peptide leader. The first, or only, vector contains the specific HRE nucleotide sequences for both oestrogen and progesterone receptors. Expression of the two chimeric proteins, results in a PDCP with two different chimeric proteins displayed. As an example, one chimeric protein could possess a binding activity for a particular ligand of interest, while the second chimeric protein could possess an enzymatic activity. Binding by the PDCP to the ligand of the first chimeric protein could then be detected by subsequent incubation with an appropriate substrate for the second chimeric protein. In an alternative embodiment a bi-functional PDCP may be created using a single DBD, by cloning one peptide at the 5'-end of the DBD, and a second peptide at the 3'-end of the DBD. Expression of this single bi-functional chimeric protein results in a PDCP with two different activities.

The applicant has investigated the possibility of screening libraries of peptides, fused to a DNA binding domain and displayed on the surface of a display package, for particular peptides with a biological activity of interest and recovering the DNA encoding that activity. Surprisingly, by manipulating the oestrogen receptor

DNA binding domain in conjunction with M13 bacteriophage the applicant has been able to construct novel particles which display large biologically functional molecules, that allows enrichment of particles with the desired specificity. The disclosure made by the applicant is important and provides a significant breakthrough in DNA library screening technology.

In order that the invention is more fully understood, embodiments will now be described in more detail by way of example only and not by way of limitation with reference to the figures below.

Description of Figures

Figure 1 shows the pDM12 N-terminal fusion oestrogen receptor DNA binding domain expression vector map, and nucleotide sequence, between the HindIII and EcoRI restriction sites, comprising a pelB leader secretion sequence, multiple cloning site containing SfiI and NotI sites, flexible $(glycine)_4$-serine linker sequence, a fragment of the oestrogen receptor gene comprising amino acids 176-282 of the full length molecule, and the 38 base pair consensus oestrogen receptor DNA binding domain HRE sequence.

Figure 2 shows the OD_{450nm} ELISA data for negative control M13K07 phage, and single-clone display particle culture supernatants (#1-4, see example 3) isolated by selection of the lymphocyte cDNA-pDM12 library against anti-human immunoglobulin kappa antibody.

Figure 3 shows partial DNA and amino acid sequence for the human immunoglobulin kappa constant region (Kabat, E.A., 1993. Sequences of proteins of immunological interest), and ELISA positive clones #2 and #3 from Figure 2 which confirms the presence of human kappa constant region DNA in-frame with the pelB leader sequence (pelB leader sequence is underlined, the cleavage site is indicated by an arrow). The differences in the 5'-end sequence demonstrates that these two clones were selected independently from the library stock. The PCR primer sequence is indicated in bold, clone #2 was amplified with CDNAPCRBAK1 and clone #3 was amplified with CDNAPCRBAK2.

Materials and Methods

The following procedures used by the present applicant are described in Sambrook, J., et al., 1989 supra.: restriction enzyme digestion, ligation, preparation of electrocompetent cells, electroporation, analysis of restriction enzyme digestion products on agarose gels, DNA purification using phenol/chloroform, preparation of 2xTY medium and plates, preparation of ampicillin, kanamycin and IPTG (Isopropyl β-D-Thiogalactopyranoside) stock solutions, preparation of phosphate buffered saline.

Restriction enzymes, T4 DNA ligase and cDNA synthesis reagents (Superscript plasmid cDNA synthesis kit) were purchased from Life Technologies Ltd (Paisley, Scotland, U.K.). Oligonucleotides were obtained from Cruachem Ltd (Glasgow, Scotland, U.K.). Taq DNA polymerase, Wizard SV plasmid DNA isolation kits, and mRNA isolation reagents (PolyATract 1000) were obtained from Promega Ltd (Southampton, Hampshire, U.K.). Kanamycin resistant M13K07 helper bacteriophage and RNAguard were obtained from Pharmacia Ltd (St. Albans, Herts, U.K.) and anti-human Igκ antibody from Seralab (Crawley Down, Sussex, U.K.).

Examples of some of the embodiments of the invention are given below in examples 1-3.

Example 1. Construction of a N-terminal PDCP display phagemid vector pDM12.

The pDM12 vector was constructed by inserting an oestrogen receptor DNA binding domain, modified by appropriate PCR primers, into a phagemid vector pDM6. The pDM6 vector is based on the pUC119 derived phage display vector pHEN1 (Hoogenboom et al., 1991, Nucleic Acids Res. 19: 4133-4137). It contains $(Gly)_4Ser$ linker, Factor Xa cleavage site, a full length gene III, and streptavidin tag peptide sequence (Schmidt, T.G. and Skerra, A., 1993, Protein Engineering 6: 109-122), all of which can be removed by NotI-EcoRI digestion and agarose gel electrophoresis, leaving a pelB leader sequence, SfiI, NcoI and PstI restriction sites upstream of the digested NotI site. The cloned DNA binding domain is under the control of the lac promoter found in pUC119.

The oestrogen receptor DNA binding domain was isolated from cDNA prepared from human bone marrow (Clontech, Palo Alto, California, U.S.A.). cDNA can be prepared by many procedures well known to those skilled in the art. As an example, the following method using a Superscript plasmid cDNA synthesis kit can be used:

(a). First strand synthesis. 5μg of bone marrow mRNA, in 5μl DEPC-treated water was thawed on ice and 2μl (50pmol) of cDNA synthesis primer (5'-AAAAGCGGCCGCACTGGCCTGAGAGA$(N)_5$-3') was added to the mRNA and the mixture heated to 70°C for 10 minutes, then snap-chilled on ice and spun briefly to collect the contents to the bottom of the tube. The following were then added to the tube:

1000u/ml RNAguard	1μl
5x first strand buffer	4μl
0.1M DTT	2μl
10mM dNTPs	1μl
200u/μl SuperScript II reverse transcriptase	5μl

The mixture was mixed by pipetting gently and incubated at 37°C for 1 hour, then placed on ice.

(b). Second strand synthesis. The following reagents were added to the first strand reaction:

DEPC-treated water	93μl
5x second strand buffer	30μl
10mM dNTPs	3μl
10u/μl *E. coli* DNA ligase	1μl
10u/μl *E. coli* DNA polymerase	4μl
2u/μl *E. coli* RNase H	1μl

The reaction was vortex mixed and incubated at 16°C for 2 hours. 2μl (10u) of T4 DNA polymerase was added and incubation continued at 16°C for 5 minutes. The reaction was placed on ice and 10μl 0.5M EDTA added, then phenol-chloroform extracted, precipitated and vacuum dried.

(c). Sal I adaptor ligation. The cDNA pellet was resuspended in 25μl DEPC-treated water, and ligation set up as follows.

cDNA	25μl
5x T4 DNA ligase buffer	10μl
1μg/μl *Sal* I adapters*	10μl

1u/μlT4 DNA ligase 5μl

*Sal I adapters: TCGACCCACGCGTCCG-3′
GGGTGCCGAGGC-5′

The ligation was mixed gently and incubated for 16 hours at 16°C, then phenol-chloroform extracted, precipitated and vacuum dried. The cDNA/adaptor pellet was resuspended in 41μl of DEPC-treated water and digested with 60 units of NotI at 37°C for 2 hours, then phenol-chloroform extracted, precipitated and vacuum dried. The cDNA pellet was re-dissolved in 100μl TEN buffer (10mM Tris pH 7.5, 0.1mM EDTA, 25mM NaCl) and size fractionated using a Sephacryl S-500 HR column to remove unligated adapters and small cDNA fragments (<400bp) according to the manufacturers instructions. Fractions were checked by agarose gel electrophoresis and fractions containing cDNA less than 400 base pairs discarded, while the remaining fractions were pooled.

(d). PCR amplification of oestrogen receptor DNA binding domain. The oestrogen receptor was PCR amplified from 5μl (150-250ng) of bone marrow cDNA using 25pmol of each of the primers pDM12FOR (5'-AAAAGAATTCTGAATGTGTTATTTTAGCTCAGGTCACTCTGACCTGATTATCAAGACCCCACTTCACCCCCT) and pDM12BAK (5'-AAAAGCGGCCGCAGGGGGAGGAGGGTCCATGGAATCTGCCAAGGAG-3') in two 50μl reactions containing 0.1mM dNTPs, 2.5 units Taq DNA polymerase, and 1x PCR reaction buffer (Promega Ltd, Southampton, U.K.). The pDM12FOR primer anneals to the 3'-end of the DNA binding domain of the oestrogen receptor and incorporates two stop codons, the 38 base pair consensus oestrogen receptor HRE sequence, and an EcoRI restriction site. The pDM12BAK primer anneals to the 5'-end of the DNA binding domain of the oestrogen receptor and incorporates the $(Gly)_4Ser$ linker and the NotI restriction site.

Reactions were overlaid with mineral oil and PCR carried out on a Techne PHC-3 thermal cycler for 30 cycles of 94°C, 1 minute; 65°C, 1 minute; 72°C, 1 minute. Reaction products were electrophoresed on an agarose gel, excised and products purified from the gel using a Geneclean II kit according to the manufacturers instructions (Bio101, La Jolla, California, U.S.A.).

(e). Restriction digestion and ligation. The PCR reaction appended NotI and EcoRI restriction sites, the $(Gly)_4Ser$ linker, stop codons and the 38 base pair oestrogen receptor target HRE nucleotide sequence to the oestrogen receptor DNA binding domain sequence (see Figure 1). The DNA PCR fragment and the target pDM6 vector (approximately 500ng) were NotI and EcoRI digested for 1 hour at 37°C, and DNA purified by agarose gel electrophoresis and extraction with Geneclean II kit (Bio101, La Jolla, California, U.S.A.). The oestrogen receptor DNA binding domain cassette was ligated into the NotI-EcoRI digested pDM6 vector overnight at 16°C, phenol/chloroform extracted and precipitated then electroporated into TG1 *E. coli* (genotype: K12, (Δlac-pro), supE, thi, hsD5/F'traD36, $proA^+B^+$, $LacI^q$, LacZΔ15) and plated onto 2xTY agar plates supplemented with 1% glucose and 100μg/ml ampicillin. Colonies were allowed to grow overnight at 37°C. Individual colonies were picked into 5ml 2xTY supplemented with 1% glucose and 100μg/ml ampicillin and grown overnight at 37°C. Double stranded phagemid DNA was isolated with a Wizard SV plasmid DNA isolation kit and the sequence confirmed with a Prism dyedeoxy cycle sequencing kit (Perkin-Elmer, Warrington, Lancashire, U.K.) using M13FOR (5'-GTAAAACGACGGCCAGT) and M13REV (5'-

GGATAACAATTTCACACAGG) oligonucleotides. The pDM12 PDCP display vector map and DNA sequence cloned between the HindIII and EcoRI restriction sites is shown in Figure 1.

Example 2. Insertion of a random-primed human lymphocyte cDNA into pDM12 and preparation of a master phage stock.

Libraries of peptides can be constructed by many methods known to those skilled in the art. The example given describes a method for constructing a peptide library from randomly primed cDNA, prepared from mRNA isolated from a partially purified cell population.

mRNA was isolated from approximately 10^9 human peripheral blood lymphocytes using a polyATract 1000 mRNA isolation kit (Promega, Southampton, UK). The cell pellet was resuspended in 4ml extraction buffer (4M guanidine thiocyanate, 25mM sodium citrate pH 7.1, 2% β-mercaptoethanol). 8ml of pre-heated (70°C) dilution buffer (6xSSC, 10mM Tris pH 7.4, 1mM EDTA, 0.25% SDS, 1% β-mercaptoethanol) was added to the homogenate and mixed thoroughly by inversion. 10µl of biotinylated oligo-dT (50 pmol/µl) was added, mixed and the mixture incubated at 70°C for 5 minutes. The lymphocyte cell lysate was transferred to 6x 2ml sterile tubes and spun at 13,000 rpm in a microcentrifuge for ten minutes at room temperature to produce a cleared lysate. During this centrifugation, streptavidin coated magnetic beads were resuspended and 6ml transferred to a sterile 50ml Falcon tube, then placed in the magnetic stand in a horizontal position until all the beads were captured. The supernatant was carefully poured off and beads resuspended in 6ml 0.5xSSC, then the capture repeated. This wash was repeated 3x, and beads resuspended in a final volume of 6ml 0.5xSSC. The cleared lysate was added to the washed beads, mixed by inversion and incubated at room temperature for 2 minutes, then beads captured in the magnetic stand in a horizontal position. The beads were resuspended gently in 2ml 0.5xSSC and transferred to a sterile 2ml screwtop tube, then captured again in the vertical position, and wash solution discarded. This wash was repeated twice more. 1ml of DEPC-treated water was added to the beads and mixed gently. The beads were again captured and the eluted mRNA transferred to a sterile tube. 50µl was electrophoresed to check the quality and quantity of mRNA, while the remainder was precipitated with 0.1 volumes 3M sodium acetate and three volumes absolute ethanol at -80°C overnight in 4 aliquots in sterile 1.5ml screwtop tubes.

Double stranded cDNA was synthesised as described in example 1 using 5µg of lymphocyte mRNA as template. cDNA was PCR amplified using oligonucleotides CDNAPCRFOR (5'-AAAGCGGCCGCACTGGCCTGAGAGA), which anneals to the cDNA synthesis oligonucleotide described in example 1 which is present at the 3'-end of all synthesised cDNA molecules incorporates a NotI restriction site, and an equimolar mixture of CDNAPCRBAK1-3

(CDNAPCRBAK1: 5'- AAAAGGCCCAGCCGGCCATGGCCCAGCCCACCACGCGTCCG,

CDNAPCRBAK2: 5'- AAAAGGCCCAGCCGGCCATGGCCCAGTCCCACCACGCGTCCG,

CDNAPCRBAK3: 5'- AAAAGGCCCAGCCGGCCATGGCCCAGTACCCACCACGCGTCCG),

all three anneal to the SalI adaptor sequence found at the 5'-end of the cDNA and incorporate a SfiI restriction site at the cDNA 5'-end. Ten PCR reactions were carried out using 2µl of cDNA (50ng) per reaction as described in example 1 using 25 cycles of 94°C, 1 minute; 60°C, 1 minute; 72°C, 2 minutes. The reactions

were pooled and a 20µl aliquot checked by agarose gel electrophoresis, the remainder was phenol/chloroform extracted and ethanol precipitated and resuspended in 100µl sterile water. 5µg of pDM12 vector DNA and lymphocyte cDNA PCR product were SfiI-NotI digested phenol/chloroform extracted and small DNA fragments removed by size selection on Chromaspin 1000 spin columns (Clontech, Palo Alto, California, U.S.A.) by centrifugation at 700g for 2 minutes at room temperature. Digested pDM12 and lymphocyte cDNA were ethanol precipitated and ligated together for 16 hours at 16°C. The ligated DNA was precipitated and electroporated in to TG1 *E. coli* and plated onto 2xTY agar plates supplemented with 1% glucose and 100µg/ml ampicillin. 10^{-4}, 10^{-5} and 10^{-6} dilutions of the electroporated bacteria were also plated to assess library size. Colonies were allowed to grow overnight at 30°C. $2x10^{8}$ ampicillin resistant colonies were recovered on the agar plates.

The bacteria were then scraped off the plates into 40ml 2xTY broth supplemented with 20% glycerol, 1% glucose and 100µg/ml ampicillin. 5ml was added to a 20ml 2xTY culture broth supplemented with 1% glucose and 100µg/ml ampicillin and infected with 10^{11} kanamycin resistance units (kru) M13K07 helper phage at 37°C for 30 minutes without shaking, then for 30 minutes with shaking at 200rpm. Infected bacteria were transferred to 200ml 2xTY broth supplemented with 25µg/ml kanamycin, 100µg/ml ampicillin, and 20µM IPTG, then incubated overnight at 37°C, shaking at 200rpm. Bacteria were pelleted at 4000rpm for 20 minutes in 50ml Falcon tubes, and 40ml 2.5M NaCl/20% PEG 6000 was added to 200ml of phage supernatant, mixed vigorously and incubated on ice for 1 hour to precipitate phage particles. Phage were pelleted at 11000rpm for 30 minutes in 250ml Oakridge tubes at 4°C in a Sorvall RC5B centrifuge, then resuspended in 2ml PBS buffer after removing all traces of PEG/NaCl with a pipette, then bacterial debris removed by a 5 minute 13500rpm spin in a microcentrifuge. The supernatent was filtered through a 0.45µM polysulfone syringe filter and stored at -20°C.

Example 3. Isolation of human immunoglobulin kappa light chains by repeated rounds of selection against anti-human kappa antibody.

For the first round of library selection a 70x11mm NUNC Maxisorp Immunotube (Life Technologies, Paisley, Scotland U.K.) was coated with 2.5ml of 10µg/ml of anti-human kappa antibody (Seralab, Crawley Down, Sussex, U.K.) in PBS for 2 hours at 37°C. The tube was rinsed three times with PBS (fill & empty) and blocked with 3ml PBS/2% BSA for 2 hours at 37°C and washed as before. $4x10^{12}$ a.r.u. of pDM12-lymphocyte cDNA phage stock was added in 2ml 2% BSA/PBS/0.05% Tween 20, and incubated for 30 minutes on a blood mixer, then for 90 minutes standing at room temperature. The tube was washed ten times with PBS/0.1% Tween 20, then a further ten times with PBS only. Bound phage were eluted in 1ml of freshly prepared 0.1M triethylamine for 10 minutes at room temperature on a blood mixer. Eluted phage were transferred to 0.5ml 1M Tris pH 7.4, vortex mixed briefly and transferred to ice. Neutralised phage were added to 10ml log phase TG1 bacteria (optical density: OD_{600nm} 0.3-0.5) and incubated at 37°C without shaking for 30 minutes, then with shaking at 200rpm for 30 minutes. 10^{-3}, 10^{-4} & 10^{-5} dilutions of the infected culture were prepared to estimate the number of phage recovered, and the remainder was spun at 4000 rpm for 10 minutes, and the pellet resuspended in 300µl 2xTY medium by vortex mixing. Bacteria were plated onto 2xTY agar plates supplemented with 1% glucose and 100µg/ml ampicillin. Colonies were allowed to grow overnight at 30°C.

A phage stock was prepared from the bacteria recovered from the first round of selection, as described in example 2 from a 100ml overnight culture. 250µl of the round 1 amplified phage stock was then selected against anti-human kappa antibody as described above with the tube was washed twelve times with PBS/0.1% Tween 20, then a further twelve times with PBS only.

To identify selected clones, eighty-eight individual clones recovered from the second round of selection were then tested by ELISA for binding to anti-human kappa antibody. Individual colonies were picked into 100µl 2xTY supplemented with100µg/ml ampicillin and 1% glucose in 96-well plates (Costar) and incubated at 37°C and shaken at 200rpm for 4 hours. 25µl of each culture, was transferred to a fresh 96-well plate, containing 25µl/well of the same medium plus 10^7 k.r.u. M13K07 kanamycin resistant helper phage and incubated at 37°C for 30 minutes without shaking, then incubated at 37°C and shaken at 200rpm for a further 30 minutes. 160µl of 2xTY supplemented with 100µg/ml ampicillin, 25µg/ml kanamycin, and 20µM IPTG was added to each well and phage amplification continued for 16 hours at 37°C while shaking at 200rpm. Bacterial cultures were spun in microtitre plate carriers at 2000g for 10 minutes at 4°C in a benchtop centrifuge to pellet bacteria and culture supernatant used for ELISA.

A Dynatech Immulon 4 ELISA plate was coated with 200ng/well anti-human kappa antibody in 100µl /well PBS for one hour at 37°C. The plate was washed 2x200µl/well PBS and blocked for 1 hour at 37°C with 200µl/well 2% BSA/PBS and then washed 2x200µl/well PBS. 50µl phage culture supernatant was added to each well containing 50µl/well 4% BSA/PBS/0.1%Tween 20, and allowed to bind for 1 hour at room temperature. The plate was washed three times with 200µl/well PBS/0.1% Tween 20, then three times with 200µl/well PBS. Bound phage were detected with 100µl/well, 1:5000 diluted anti-M13-HRP conjugate (Pharmacia) in 2% BSA/PBS/0.05% Tween 20 for 1 hour at room temperature and the plate washed six times as above. The plate was developed for 5 minutes at room temperature with 100µl/well freshly prepared TMB (3,3',5,5'-Tetramethylbenzidine) substrate buffer (0.005% H_2O_2, 0.1mg/ml TMB in 24mM citric acid/52mM sodium phosphate buffer pH 5.2). The reaction was stopped with 100µl/well 12.5% H_2SO_4 and read at 450nm. ELISA (ELISA data for binding clones is shown in Figure 2). These clones were then sequenced with M13REV primer as in example 1. The sequence of two of the clones isolated is shown in Figure 3.

Claims

1. A method of screening a DNA library for nucleotide sequences which encode a peptide of interest, by:

(a). Constructing a library comprising at least 10^5 different peptide display carrier packages, each package comprising:

(i) a chimeric protein comprising (1) a peptide binding domain, which is stable in structure, and which binds to a specific target, and (2) at least a functional portion of a DNA binding protein, said protein acting, when the chimeric protein is produced under suitable conditions in a suitable host cell, to link the chimeric binding protein or a processed form thereof, to the DNA encoding the chimeric protein.

(ii) a protective coat supplied by a replicable carrier package capable of independent existence, which is not required for linking the displayed peptide to the DNA encoding that peptide but allows the display of the expressed peptide binding domain component of the chimeric protein on the outer surface of the protective coat, and provides a mechanism for releasing the intact PDCP from a host cell.

(b). Selecting a PDCP encoding the peptide fragment that binds specifically to the target of interest by binding said particle to the target specific for said peptide fragment, removing the library particles that do not bind to said target, amplifying the selected PDCP's in suitable host cells, and repeating the selection against the target of interest and amplification of binding particles, until individual binding PDCP's can be detected in a suitable binding assay.

2. A nucleotide sequence of the chimeric fusion protein of claim 1 that consists of a peptide binding domain fused to a DNA binding domain by a flexible linker peptide in a phagemid DNA vector.
3. A replicable carrier package of claim 1 that is a filamentous bacteriophage.
4. A phagemid vector of claim 2 where the DNA binding domain is a member of the nuclear hormone receptor family of DNA binding proteins, including the oestrogen receptor, or fragments of said proteins.
5. A vector of claim 4 containing one or more copies of the specific nucleotide sequence, bound by the DNA binding protein.
6. A library of claim 1 where said peptide binding domain is encoded by:

(a). a cDNA fragment isolated from a cell type of interest.

(b). a single-chain antibody comprising a heavy chain variable region linked to a light chain variable region by a flexible amino acid linker.

(c). a heavy chain antibody variable region and a heavy chain constant region domain.

(d). a light chain antibody variable region and the light chain constant region domain.

(e). a synthetic DNA fragment.

Chimeric binding peptide library screening method

Abstract

Nucleotide sequences encoding binding peptides are isolated from DNA libraries using DNA binding proteins to link the peptide to the sequence which encodes it. DNA libraries are prepared from cells encoding the protein of interest, or from synthetic DNA, and inserted into, or adjacent to, a DNA binding protein in an expression vector to create a chimeric fusion protein. Incorporation of the vector DNA into a carrier package, during expression of the chimeric fusion protein, results in the production of a peptide display carrier package (PDCP) displaying the DNA-bound fusion protein on the external surface of the carrier package. Employment of affinity purification techniques results in the PDCP particles containing sequences encoding the desired peptide to be selected and the desired nucleotide sequences obtained therefrom.

Figure 1

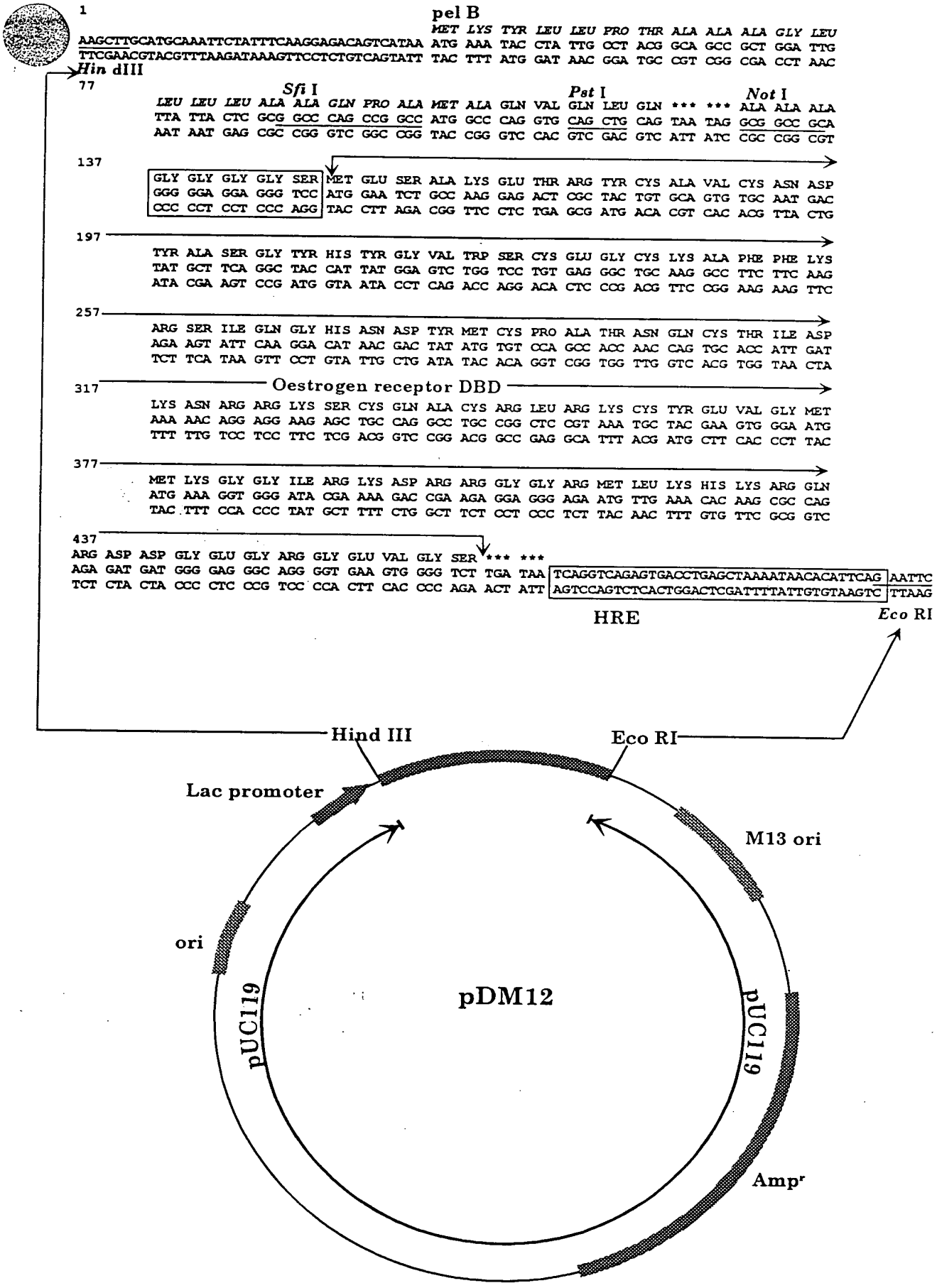
pel B
Hin dIII
Sfi I
Pst I
Not I
Oestrogen receptor DBD
HRE
Eco RI
Hind III
Eco RI
Lac promoter
M13 ori
ori
pUC119
pDM12
pUC119
Ampr

Figure 2

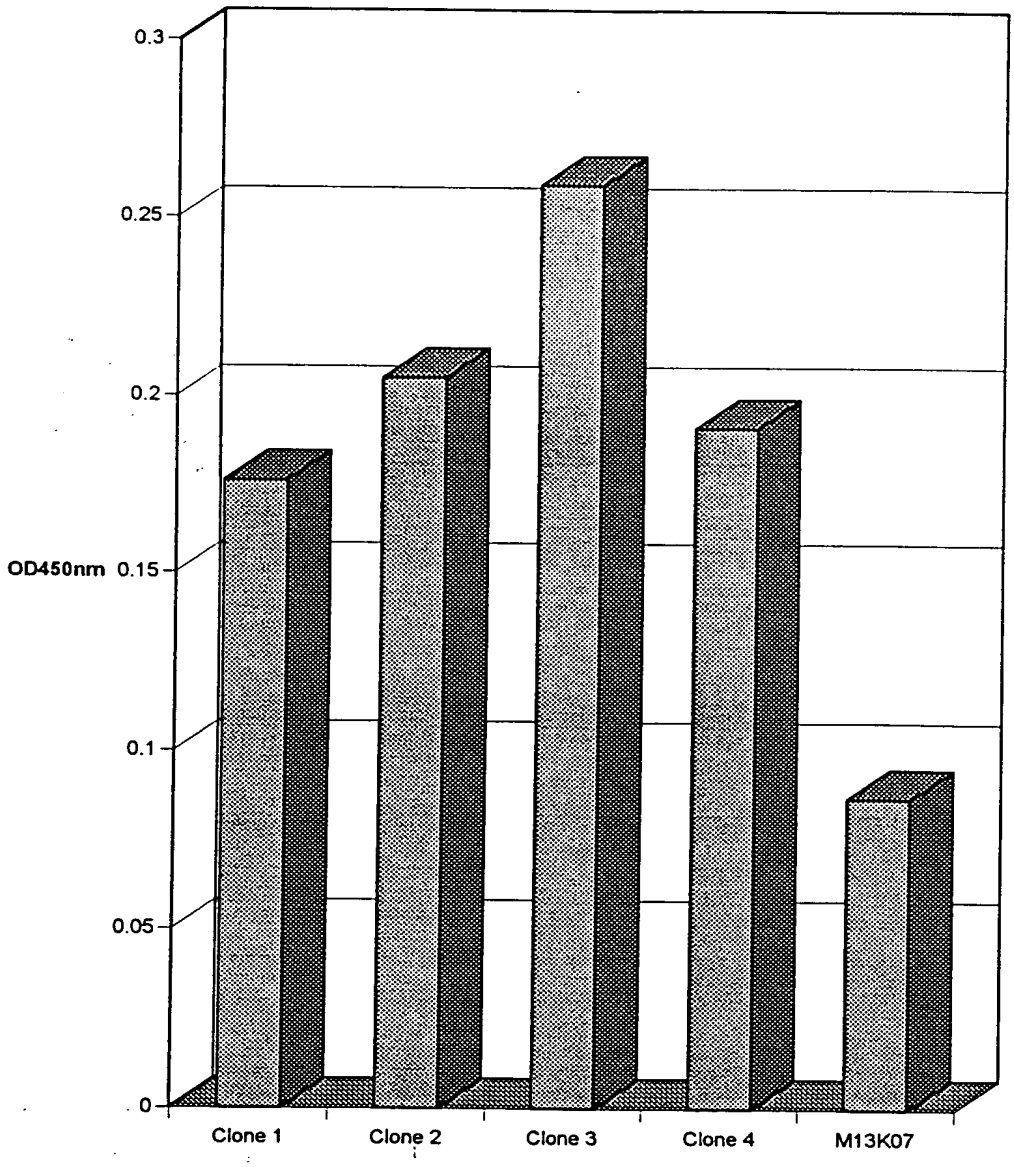
0.3
0.25
0.2
OD450nm 0.15
0.1
0.05
0
Clone 1
Clone 2
Clone 3
Clone 4
M13K07

Figure 3

Human Igκ constant region

K R T V A A P S V
AAACGAACTGTGGCTGCACCATCTGTC

Clone #2

M A ↓Q P T T R P G Q G T R L D I K R T V A A P S V
ATGGCCCAGCCCACCACGCGTCCGGGCCAAGGGACACGACTGGACATTAAACGAACTGTGGCTGCACCATCTGTC

Clone #3

M A ↓Q S H H A S G G G T K V E I K R T V A A P S V
ATGGCCCAGTCCCACCACGCGTCCGGCGGAGGGACCAAGGTGGAGATCAAACGAACTGTGGCTGCACCATCTGTC

Human Igκ constant region

F I F P P S D E Q L K S G T A S V V C L L N N F Y
TTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTAT

Clone #2

F I F P P S D E Q L K S G T A S V V C L L N N F Y
TTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTAT

Clone #3

F I F P P S D E Q L K S G T A S V V C L L N N F Y
TTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTAT

PCT / GB 92 / 02[illegible]

Margit[illegible]

21/9/92